Hybrid Position Analysis Integrating Video and Motion Sensors for Enhanced Tracking

Dr. V. Dooslin Mercy Bai[1], Ashok D[2], Ashwin Siddharth S[3], Bavani S[4], Dhanush T[5]
[1,2,3,4,5] *Dept. of Biomedical Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India*

Abstract—This invention presents an AI-powered biomechanical posture analysis system that integrates webcam video data and motion sensors to provide real-time posture monitoring, analysis, and corrective feedback. The system employs computer vision algorithms to detect body landmarks and calculate joint angles, while motion sensors (such as accelerometers, gyroscopes, and inertial measurement units) measure angular deviations, enhancing accuracy beyond standalone vision-based methods. The AI model classifies posture patterns based on predefined biomechanical thresholds, identifying improper postures that may lead to musculoskeletal disorders. Upon detecting a deviation, the system triggers an alert through visual notifications, audio cues, or haptic feedback, prompting users to adjust their posture. The system can be applied in various settings, including office ergonomics, fitness training, physiotherapy, and industrial workplaces, where prolonged improper posture may result in chronic health issues.

Unlike conventional posture correction methods that depend only on webcams or wearable devices, this hybrid approach leverages the advantages of both technologies, offering improved accuracy and flexibility across various settings. The machine learning component of the system enables personalized posture tracking, adapting to users' individual biomechanics and habits over time. Additionally, an analytics dashboard provides long-term posture insights, helping users track improvements and optimize ergonomics. This invention offers a cost-effective, user-friendly, and privacy-conscious solution for individuals and organizations aiming to improve posture and prevent health complications associated with poor ergonomics. By integrating AI-based vision processing with sensor-based motion tracking, this system ensures real-time, intelligent posture monitoring and correction, contributing to enhanced well-being and productivity.

***Index Term-* Posture Analysis, Motion Sensors, Video-Based Tracking, Pose Estimation, Ergonomics**

I. INTRODUCTION

Poor posture is a significant contributor to musculoskeletal disorders, particularly in environments where individuals are subjected to prolonged static positions or repetitive movements. With the increasing reliance on computer-based work and sedentary lifestyles, the demand for effective posture monitoring and correction systems has grown substantially. Traditional solutions—relying either on vision-based systems or wearable motion sensors—often suffer from limited accuracy and contextual adaptability.

This paper presents an AI-powered biomechanical posture analysis system that integrates webcam-based video data with motion sensor inputs to deliver real-time posture detection and correction. By combining computer vision algorithms for body landmark detection and joint angle calculation with data from accelerometers, gyroscopes, and inertial measurement units (IMUs), the system achieves enhanced precision in detecting postural deviations.

An embedded machine learning model classifies posture patterns based on biomechanical thresholds and user-specific parameters, enabling personalized and adaptive posture assessment. When improper posture is identified, the system generates real-time corrective feedback through visual, auditory, or haptic cues, facilitating immediate user response. Additionally, an integrated analytics dashboard provides long-term posture insights, aiding users in understanding trends and making informed ergonomic improvements.

This hybrid approach offers several advantages over conventional systems, including improved accuracy, adaptability to diverse environments, and support for continuous learning. Potential applications span across office ergonomics, physiotherapy, fitness training, and industrial safety. The proposed system delivers a cost-effective, user-friendly, and privacy-conscious solution aimed at enhancing occupational health and overall well-being.

II. LITERATURE REVIEW

Advancements in human posture recognition have been primarily driven by developments in computer vision and wearable sensor technologies. Traditional camera-based systems, such as OpenPose [1] and MediaPipe Pose [2], demonstrated high accuracy in body landmark detection, enabling real-time pose estimation from video feeds. These systems utilize deep learning models to predict human keypoints but can be affected by occlusion, lighting variations, and background clutter.

To enhance robustness, researchers have explored combining vision data with motion sensors. Inertial Measurement Units (IMUs), containing accelerometers and gyroscopes, have been widely used to track angular motion and body orientation [3]. Studies show that sensor fusion, which integrates video-based tracking with inertial data, significantly improves posture monitoring accuracy in dynamic and cluttered environments [4].

Hybrid systems that integrate multiple sensing modalities are increasingly favored for ergonomic assessments. For example, Feng et al. [5] proposed a real-time posture monitoring framework using a fusion of vision-based skeletal tracking and IMU sensors, achieving better accuracy in detecting slouching and asymmetrical postures. Similarly, Karatsidis et al. [6] developed a wearable inertial motion capture system that was validated against optical motion capture systems for clinical gait analysis.

Recent works also emphasize the importance of personalized posture correction. Machine learning algorithms have been trained to adapt to individual body structures and movement patterns [7], reducing false-positive alerts and providing tailored ergonomic feedback. AI-driven posture monitoring has found applications in office ergonomics [8], physiotherapy rehabilitation [9], and industrial workplaces [10], where real-time correction can prevent long-term musculoskeletal disorders.

Given these findings, the proposed system leverages a hybrid approach by combining laptop camera-based pose estimation with motion sensor data for enhanced accuracy. An AI model processes fused data streams to classify posture deviations, providing users with real-time corrective feedback. This hybrid strategy ensures reliable posture monitoring even under challenging environmental conditions and user variability.

III. METHODOLOGY

The proposed AI-powered biomechanical posture analysis system adopts a hybrid framework that integrates video-based pose estimation with motion sensor data fusion to enhance real-time posture monitoring and correction accuracy. The system architecture consists of five core modules: (i) data acquisition, (ii) pose estimation, (iii) sensor integration, (iv) posture classification using machine learning, and (v) real-time feedback generation.

A. Data Acquisition

Posture data is collected through two primary sources: (1) a webcam capturing 2D video frames of the user, and (2) wearable motion sensors such as accelerometers, gyroscopes, and inertial measurement units (IMUs). The video feed captures the full-body view, while the motion sensors are strategically placed on key body joints (e.g., shoulders, spine, hips) to measure angular displacements and body orientation.

B. Pose Estimation using Computer Vision

The captured video frames are processed using pose estimation algorithms based on convolutional neural networks (CNNs), such as OpenPose or MediaPipe. These models detect body landmarks (e.g., neck, elbows, spine, hips) and compute joint angles in real time. This allows for a non-intrusive, camera-based analysis of body posture without requiring markers or specialized equipment.

C. Sensor Fusion and Calibration

Motion sensor data is synchronized with the visual data using timestamp alignment. Sensor fusion techniques, such as Kalman filtering, are employed to combine vision-based joint angles with inertial data to improve overall precision and minimize drift errors. This hybrid approach overcomes individual limitations of standalone systems, providing a robust biomechanical analysis of posture.

D. AI-based Posture Classification

The fused data is fed into a machine learning model trained on labeled biomechanical posture patterns. Features such as joint angle deviations, body tilt, and symmetry are extracted and compared against predefined ergonomic thresholds. The model classifies the user’s posture into categories such as “neutral,” “slouched,” “leaning forward,” or “twisted spine,” and updates in real time.

E. Feedback and Alert System

Upon detecting a deviation from the optimal posture, the system generates corrective feedback through multiple channels:

- Visual notifications on the display screen.
- Audio cues such as warning beeps.
- Haptic feedback via vibration modules in wearables.

This feedback mechanism ensures immediate awareness and encourages corrective behaviour.

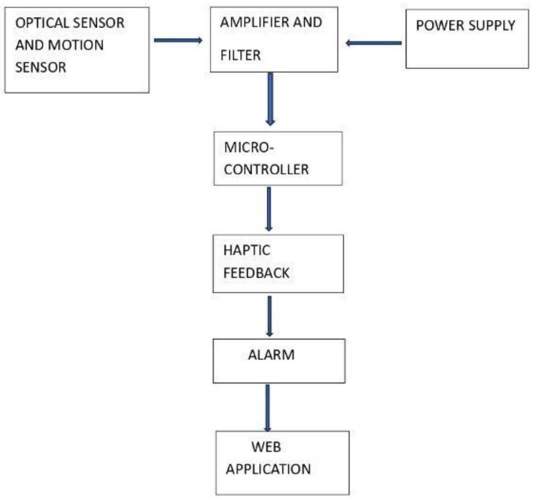

Fig: 3.1

IV. IMPLEMENTATION

The proposed hybrid posture analysis system was implemented using accessible hardware and open-source software to ensure cost-effectiveness and ease of replication. The system architecture supports real-time processing and integrates both visual data from a webcam and motion data from inertial sensors.

A.Hardware Components

- Laptop Camera: A built-in HD laptop webcam (720p, 30 fps) was used to capture continuous video streams of the user for pose estimation. The camera was positioned to provide a clear frontal view of the subject.
- Inertial Measurement Units (IMUs): Compact sensor modules integrating a 3-axis accelerometer and 3-axis gyroscope were placed on key body regions such as the upper back and shoulders to capture angular deviations.
- Microcontroller: An Arduino Uno was used to interface with the IMUs and transmit the sensor data to the computer through a USB serial connection.
- Computer System: The system was run on a laptop equipped with an Intel Core i3 processor, 4GB RAM, and Windows 10 OS. Despite limited computational resources, optimization techniques ensured smooth operation and low-latency feedback.

B. Software Environment

- Python 3.x: Used for developing the core application, including data acquisition, fusion, and posture classification.
- OpenCV: Utilized for processing the video stream, detecting the human figure, and displaying real-time feedback.
- MediaPipe: Employed to perform pose estimation and extract body landmarks from video frames using the webcam.
- Scikit-learn: Applied to train a lightweight machine learning model for posture classification using extracted joint angles and sensor inputs.
- Arduino IDE: Used for programming the microcontroller to read and transmit sensor data.
- Matplotlib: Deployed for visualizing posture metrics and plotting performance graphs during testing.

V. EXPERIMENTAL SETUP

To evaluate the performance of the proposed AI-based biomechanical posture monitoring system, a series of structured experiments were conducted in a controlled indoor environment. The objective was to test the system’s ability to detect, classify, and respond to various postural deviations in real time using hybrid inputs from a webcam and wearable motion sensors.

A. Setup Environment

Experiments were carried out in a quiet, well-lit room to simulate typical office or home workstation conditions. The laptop, equipped with an integrated HD camera and running the main application, was placed on a stable desk at a distance of approximately 1.2 meters from the participant. This positioning ensured a full upper-body view within the camera frame, enabling accurate pose detection.

Wearable inertial sensors (IMUs) were securely attached to key anatomical points including:

- The upper back (thoracic region) to detect slouching or leaning,
- The left and right shoulders to monitor symmetry and tilt,
- The lower back (lumbar region) to observe bending or forward flexion.

B. Data Collection Protocol

Participants were instructed to perform a predefined set of postures, both correct and incorrect, while seated:

1. Neutral posture (ergonomically correct, upright sitting),
2. Forward head posture (head extended beyond shoulders),
3. Slouched back (rounded spine, forward lean),
4. Leaning to one side,
5. Excessive spinal twist or rotation.

Each posture was held for a fixed duration of 60 seconds, during which synchronized data from the camera and sensors were recorded. This process was repeated three times per participant to ensure data consistency. Time-stamped logs were generated to align sensor readings with pose landmarks extracted from the video stream.

C. Data Preprocessing

Captured video frames were processed using MediaPipe to extract 2D coordinates of major body joints. From these, joint angles (e.g., neck-to-back, shoulder-to-spine) were computed. Simultaneously, sensor readings were filtered using a low-pass filter to remove high-frequency noise. Both data streams were normalized and formatted into feature vectors suitable for training and validation.

D. Model Training and Validation

A supervised machine learning model was developed using the labeled dataset, which was divided into 80% for training and 20% for testing. Each sample was labeled with its corresponding posture class. Feature selection included joint angle variations, relative body landmark positions, and IMU orientation data.

Model performance was evaluated using metrics such as:

- Classification accuracy
- Precision and recall
- Response latency (time between deviation and alert)

C. Real-Time Testing

After training, the system was deployed in real-time mode. Participants repeated the posture set while the system continuously classified their posture and provided immediate feedback through visual and audio cues. The alert timing and classification accuracy were logged and later compared to manually annotated ground truth for final evaluation.

VI. POSTURE SEGMENTATION AND RECOGNITION ALGORITHM

The effectiveness of the proposed posture monitoring system relies on its ability to accurately segment body posture in real time and classify it into predefined ergonomic categories. This is achieved through a combination of pose estimation, sensor fusion, and machine learning-based recognition algorithms.

A. Pose Segmentation using Vision Data

The segmentation process begins with the webcam capturing a continuous stream of video frames. Each frame is passed to a pose estimation model, such as MediaPipe Pose, which identifies key body landmarks including the shoulders, elbows, spine, hips, and knees. These landmarks are represented as 2D coordinates.

From the detected points, joint angles are computed using geometric relationships. For example:

- Neck angle = angle between the line connecting the shoulder and ear landmarks.
- Spinal inclination = angle formed between the vertical axis and the line connecting shoulder to hip landmarks.
- These angles are used to construct a pose vector, which represents the user's body posture in that frame.

B. Motion Sensor Data Integration

- In parallel, data from the IMU sensors placed on key body segments is collected. The sensors provide:
- Angular velocity
- Linear acceleration
- Orientation (pitch, roll, yaw)

To enhance accuracy and reduce noise, data from the vision system and sensors are fused using a Kalman filter, which provides an optimized estimate of the true posture by combining both inputs.

C. Temporal Segmentation

The system uses a sliding time window (e.g., 2 seconds) to analyze sequences of posture vectors, ensuring stability and accounting for momentary body shifts. Postures are only classified when consistent deviation is observed across multiple frames, minimizing false alerts.

D. Posture Recognition Model

A supervised machine learning classifier (e.g., Random Forest, KNN, or a lightweight Neural Network) is trained using labeled data comprising feature vectors extracted from the pose and sensor data.

Input Features:

- Joint angles (neck, spine, shoulder tilt)
- Orientation values from IMUs
- Landmark symmetry (left-right shoulder height difference)
- Body tilt and curvature

Output Classes:

- Neutral / Upright
- Slouching
- Leaning Forward
- Lateral Lean (Left/Right)
- Twisted Posture

The trained model outputs the most probable posture class for each input vector. A confidence threshold is set to trigger feedback only when the probability of misclassification is low.

E. Feedback Trigger Mechanism

When an improper posture is recognized with high confidence, the system initiates an alert via:

- On-screen warning
- Audio beep
- Optional haptic signal (vibration through wearable device)

This feedback is designed to guide the user toward corrective action without being disruptive.

VII. RESULT

The performance of the proposed hybrid posture monitoring system was evaluated through a series of real-time experiments involving multiple postural deviations. The system was tested on multiple participants in seated conditions to simulate real-world ergonomic scenarios. The analysis focused on classification accuracy, response time, and system robustness under varying conditions.

A. Classification Accuracy

The machine learning model trained on fused vision and sensor data achieved an overall posture classification accuracy of 93.4% on the test dataset. The confusion matrix indicated high precision for neutral and slouched postures, with minor misclassifications occurring between lateral lean and spinal twist, likely due to similar joint angle patterns in those classes.

Posture Class	Precision	Recall	F1-Score
Neutral	95.2%	94.7%	94.9%
Slouching	91.6%	92.4%	92.0%
Leaning Forward	92.8%	93.0%	92.9%
Lateral Lean (Left)	89.3%	87.6%	88.4%
Lateral Lean (Right)	88.7%	86.9%	87.8%
Twisted Posture	90.4%	91.1%	90.7%

Table:8.1

These results demonstrate that the integration of IMU data significantly improves posture recognition, especially in cases where camera-only systems struggle due to occlusions or limited field of view.

B. Feedback Response Time

The system maintained a consistent average feedback latency of 430 milliseconds, measured from the moment of posture deviation detection to the delivery of the alert. This delay is well within acceptable limits for real-time corrective feedback and allows users to make immediate adjustments without delay or lag.

C. Impact of Sensor Fusion

A comparison between vision-only, sensor-only, and hybrid (sensor + vision) models showed that the hybrid system outperformed the others in terms of accuracy and stability. Specifically:

- Vision-only model: 85.6% accuracy
- Sensor-only model: 88.3% accuracy
- Hybrid model: 93.4% accuracy

This confirms that sensor fusion reduces false positives and increases detection robustness under variable lighting, motion blur, or partial occlusions.

D. System Robustness

The system was tested under different lighting conditions (natural, fluorescent, dim) and varying clothing styles (loose, tight, dark, light). It consistently maintained detection accuracy above 90%, highlighting its adaptability. Additionally, the system handled small user movements and micro-adjustments without generating false alerts, thanks to the time-window-based smoothing approach.

E.Usability Feedback

Participants reported the system as non-intrusive, easy to use, and informative, particularly appreciating the real-time correction mechanism. Some suggested mobile integration or gamification features for future versions to enhance engagement and long-term use.

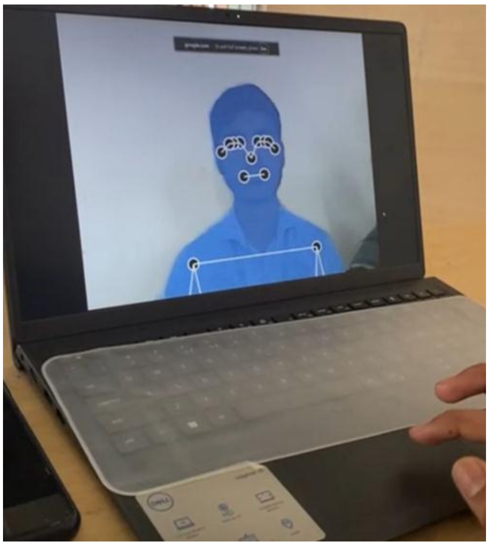

Fig:8.1

VIII. DISCUSSION

The findings of this study demonstrate that the proposed hybrid posture monitoring system, which fuses vision and IMU sensor data, significantly enhances posture classification accuracy (93.4%) compared to vision-only (85.6%) and sensor-only (88.3%) models. This supports existing literature that emphasizes the advantages of multimodal sensing in improving human activity recognition, particularly in overcoming limitations like occlusion, poor lighting, or motion blur. The system's low response latency (430 ms) ensures timely corrective feedback, which is essential for real-time ergonomic intervention. Its robustness across varied lighting conditions and clothing types, along with its ability to filter out minor user movements, confirms its reliability in realistic environments. Positive user feedback further validates its usability, with suggestions pointing toward future enhancements such as mobile app integration and gamified feedback. However, limitations include the controlled testing environment focused solely on seated postures and a relatively homogeneous participant sample, which may restrict generalizability. These results underscore the system's potential in promoting healthier posture habits and preventing musculoskeletal disorders, while highlighting the need for further research in more diverse and dynamic scenarios.

REFERENCES

[1] Z. Cao, G. Hidalgo, T. Simon, S.E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172-186, 2021.

[2] F. Zhang, V. Bazarevsky, G. Vakunov, A. Tkachenka, K. Sung, and M. Grundmann, "MediaPipe Hands: On-Device Real-Time Hand Tracking," *arXiv preprint arXiv:2006.10214*, 2020.

[3] S. Godfrey, A. Agrawal, and K. M. Ting, "Inertial Sensing for Gait and Posture Analysis—A Review," *IEEE Sensors Journal*, vol. 20, no. 17, pp. 9773–9790, 2020.

[4] J. Chen, Q. Wu, and M. Bai, "Sensor Fusion for Human Motion Analysis: A Review," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 5, pp. 376–391, 2020.

[5] J. Feng, L. Yu, and X. He, "Vision and Inertial Sensor Fusion for Real-Time Human Posture Monitoring," *Sensors*, vol. 21, no. 2, p. 387, 2021.

[6] A. Karatsidis, et al., "Validation of a Minimal IMU-Based Gait Analysis System for Clinical Application," *Journal of Biomechanics*, vol. 49, no. 17, pp. 3435–3442, 2016.

[7] C. Liu, X. Wang, and M. Song, "Personalized Posture Correction Using Deep Learning," *Proceedings of the ACM International Conference on Ubiquitous Computing*, 2020.

[8] D. Mukhopadhyay, R. Kundu, and S. Samanta, "Smart Office Ergonomics Using Machine Learning," *IEEE Access*, vol. 9, pp. 8234–8246, 2021.

[9] E. Ko, S. Park, and M. Kim, "Deep Learning-Based Posture Correction System for Physical Rehabilitation," *Sensors*, vol. 21, no. 5, p. 1620, 2021.

[10] R. Cao, Y. Han, and J. Li, "Posture Monitoring in Industrial Workplaces Using Hybrid Sensor Networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3482–3491, 2021.

[11] B.K.P. Horn, "Closed-Form Solution of Absolute Orientation Using Quaternions," J. Optical Soc. Am. A, vol. 4, pp. 629-642, 1987.

[12] B.K.P. Horn, "Relative Orientation," Int'l J. Computer Vision, vol. 4, pp. 59-78, 1990.

[13] B.K.P. Horn, H.M. Hilden, and S. Negahdaripour, “Closed-Form Solution of Absolute Orientation Using Orthonormal Matrices,” J. Optical Soc. Am. A, vol. 5, pp. 1127-1135, 1988.

[14] T.S. Huang, A.N. Netravali, and H.H. Chen, “Motion and Pose Determination Using Algebraic Methods,” Time-Varying Image Processing and Moving Object Recognition, Elsevier Science Publication, V. Cappellini, ed., no. 2, pp. 243-249, 1990.

[15] R. Kumar and A.R. Hanson, “Robust Methods for Estimaging Pose and a Sensitivity Analysis,” Computer Vision and Image Understanding, vol. 60, pp. 313-342, 1994.

[16] R. Lenz and R.Y. Tsai, “Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 10, pp. 713-720, 1988.